STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Production, purification and crystallization of a *trans*-sialidase from *Trypanosoma vivax*

Carole L. F. Haynes,[a,b,c] Paul Ameloot,[c] Han Remaut,[a,d] Nico Callewaert,[c] Yann G.-J. Sterckx[a,b]*‡ and Stefan Magez[a,b]‡

[a]Structural Biology Research Center (SBRC), VIB, Pleinlaan 2, B-1050 Brussels, Belgium, [b]Research Unit for Cellular and Molecular Immunology (CMIM), VUB, Pleinlaan 2, B-1050 Brussels, Belgium, [c]Department for Molecular Biomedical Research (DMBR), UGent, Ghent, Belgium, and [d]Structural and Molecular Microbiology (SMM), VUB, Pleinlaan 2, B-1050 Brussels, Belgium. *Correspondence e-mail: yann.sterckx@vib-vub.be

Received 26 December 2014
Accepted 5 February 2015

Edited by W. N. Hunter, University of Dundee, Scotland

‡ These authors contributed equally to this work and should be considered joint senior authors.

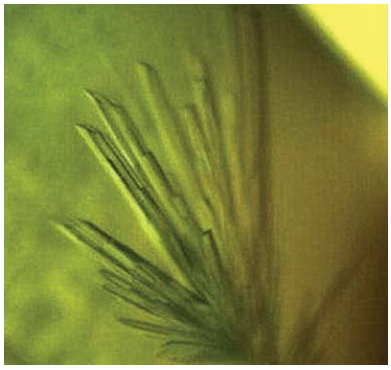
Keywords: *trans*-sialidase; *Trypanosoma vivax*.

Sialidases and *trans*-sialidases play important roles in the life cycles of various microorganisms. These enzymes can serve nutritional purposes, act as virulence factors or mediate cellular interactions (cell evasion and invasion). In the case of the protozoan parasite *Trypanosoma vivax*, *trans*-sialidase activity has been suggested to be involved in infection-associated anaemia, which is the major pathology in the disease nagana. The physiological role of trypanosomal *trans*-sialidases in host–parasite interaction as well as their structures remain obscure. Here, the production, purification and crystallization of a recombinant version of *T. vivax trans*-sialidase 1 (rTvTS1) are described. The obtained rTvTS1 crystals diffracted to a resolution of 2.5 Å and belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 57.3, b = 78.4, c = 209.0 Å.

1. Introduction

Sialic acids (SA) are negatively charged sugars that are present in a multitude of complex carbohydrates on the surfaces of mammalian cells, including red blood cells (RBCs; Chen & Varki, 2010). These sugars play important roles in a myriad of biological processes, such as cell adhesion, intercellular communication and self-recognition processes, regulation of immune responses and apoptosis (Engstler *et al.*, 1993; Pshezhetsky & Ashmarina, 2013). Interestingly, SA are also key components for host–pathogen interactions. Like the cells of the animal host, pathogens and commensals can decorate their surfaces with SA (Varki & Gagneux, 2012). This helps pathogens to modulate the host's immune system, to avoid immune recognition and to invade host cells (Stafford *et al.*, 2012; Montagna *et al.*, 2006; Schenkman *et al.*, 1991; Blom-Potar *et al.*, 2010; Muiá *et al.*, 2010; Schauer & Kamerling, 2011). In addition, some pathogens require host-derived SA for nutritional requirements. While mammals can synthesize SA *de novo via* a complex biochemical pathway, it seems that many host-associated viral, fungal, bacterial and parasitic pathogens partially or completely lack the genes necessary for SA biosynthesis. Instead, these pathogens have evolved to scavenge SA from their host organism.

One of the most common strategies that SA auxotroph pathogens employ to procure host SA is through the action of pathogen-encoded sialidases and *trans*-sialidases. Sialidases are able to cleave off SA from host sialyl donors and release the product into the surrounding environment. Even though the overall sequences of sialidases are divergent, a so-called 'sialidase motif' and the active-site residues are highly



conserved. Enzymatic removal of SA from the host cells by means of a pathogen's sialidase(s) results in exposure of the underlying glycans, thus perturbing natural self-recognition phenomena and potentially increasing inflammatory responses by exposing desialylated danger-associated molecular patterns (DAMPS; Chen & Nuñez, 2010; Varki, 2011). Moreover, diminished SA surface levels reduce the surface hydrophilicity of host cells, decrease charge repulsion between host cells and disturb the endogenous receptor interactions. The damaging effects of SA disturbances could explain why active host sialidases are not present in extracellular fluids. The closely related *trans*-sialidases are also used by a subset of pathogens to scavenge SA from the host organism. *Trans*-sialidase enzymes share the 'sialidase motif' but have a distinct mode of action. The members of this enzyme class use host glycoconjugates as sialyl donors, but transfer the SA to parasite acceptor molecules rather than water. Coinciding with their highly conserved sialidase-like active site, it has been shown that *trans*-sialidases exhibit residual sialidase activity in the absence of the appropriate acceptor (Cross & Takle, 1993; Harrison *et al.*, 2001; Gbem *et al.*, 2013). In contrast to sialidases, it is interesting to note that the structural and mechanistic features of *trans*-sialidases remain understudied, as only a handful of these enzymes have been characterized (Cheng *et al.*, 2008, 2009; Gut *et al.*, 2008; Newstead *et al.*, 2004).

The members of the *Trypanosoma* genus are protozoan parasites that infect and can cause disease in cattle and humans, mainly in Africa and South America. In the case of cattle infection, the induced disease is called nagana, and is caused mainly by *T. congolense* and *T. vivax*, and to a lesser extent by *T. brucei*. To successfully complete their life cycle, the African cattle-infecting trypanosomes (*T. brucei* and *T. congolense*) require alternate passages through insects of the *Glossina* species and a mammalian host. In contrast, *T. vivax* does not depend on the insect-vector stage to complete its life cycle. Instead, *T. vivax* has evolved to employ any species of biting fly for mechanical transmission into a new host. This has allowed *T. vivax* to spread further to South America through cattle transport. Given the severity of the consequences of trypanosome infection in cattle (weight loss, diminished milk production and death; Taylor, 1998) and the inability to treat and/or clear the infection (Barrett & Fairlamb, 1999; Taylor & Mertens, 1999; La Greca & Magez, 2011), nagana causes severe economic losses. Like many other host-associated pathogens, trypanosomes lack genes for SA biosynthesis and scavenge SA from their host by employing sialidases and *trans*-sialidases. In some trypanosomes (for example, the pathogenic *T. brucei* and the nonpathogenic *T. rangeli*), the use of their (*trans*-)sialidases is restricted to the insect stage and is believed to provide protection of the parasite during its journey through the insect midgut (Montagna *et al.*, 2002). In other cases, trypanosomal (*trans*-) sialidases are produced in the so-called 'bloodstream form' (*i.e.* the mammalian host stage) and their actions are thought to be directly correlated with the induced pathology in the infected mammalian host (Schenkman *et al.*, 1991; Nok & Balogun, 2003; Buratai *et al.*, 2006). In *T. cruzi*, the only intracellular trypanosome, it has been shown that its *trans*-sialidases play a role in host-cell invasion, parasite virulence and immune suppression (Schenkman *et al.*, 1992; Buschiazzo *et al.*, 2002; Jacobs *et al.*, 2010; Giorgi & de Lederkremer, 2011). In *T. vivax* infections, *trans*-sialidase activity has been proposed to be linked to the typically observed severe anaemia (Bratosin *et al.*, 1998; Buratai *et al.*, 2006; Guegan *et al.*, 2013). Hence, for disease-associated trypanosomal *trans*-sialidases, unravelling their structure–function relationship might yield insights into the molecular mechanisms behind the onset of trypanosomosis-induced pathologies.

In this article, we describe the production, purification, crystallization conditions and preliminary X-ray analysis of the recombinant *T. vivax trans*-sialidase 1 (rTvTS1). Studying this enzyme is of interest because of its low sequence identity to other trypanosomal sialidases and *trans*-sialidases with known structures (30.1% to that from *T. rangeli* and 24.9% to that from *T. cruzi*), its distinct biological function and its possible medical relevance (Buschiazzo *et al.*, 2002; Amaya *et al.*, 2003, 2004). It is expected that crystal structure determination of rTvTS1 will not only contribute to increasing the knowledge of the understudied class of *trans*-sialidase enzymes but will also provide insights into its biological role in *T. vivax* infections.

2. Materials and methods

2.1. Construction of the expression vectors and transfection in *Pichia pastoris*

Three expression vectors (named pPpT4_αS_TvTS1, pPpT4_αS_TvTS2 and pPpT4_αS_TvTS3) were constructed and respectively encode three representative *trans*-sialidase (TS) variants from *T. vivax* Y486 (TvTS1, GenBank ID CAEX01006974; TvTS3, GenBank ID CAEX01004555; TvTS4, GenBank ID HE573024). The native coding

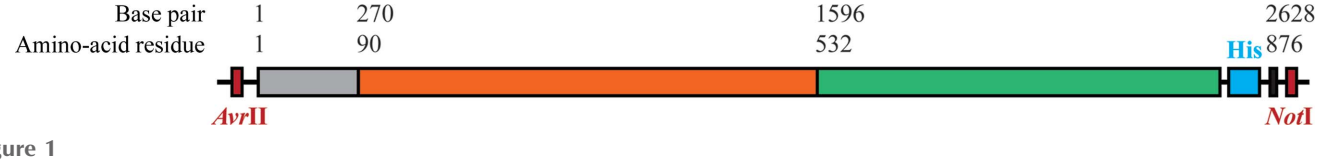


Figure 1
Schematic overview of the synthesized rTvTS1-encoding gene. The catalytic and lectin-like domains are indicated by orange and green boxes, respectively. The C-terminal hexahistidine tag is displayed as a blue box and the grey box shows a domain of unknown function. The AvrII and NotI restriction sites necessary for subcloning are also indicated (red boxes). The resulting protein product has a length of 876 residues and a theoretical molecular mass of 97.6 kDa.

sequences were engineered to facilitate protein production in and secretion by the methylotrophic yeast *P. pastoris* (Fig. 1). Firstly, the predicted signal peptides of the *trans*-sialidases (except for TvTS3, which does not contain such a leader peptide) were removed. Secondly, AvrII and NotI restriction sites were introduced for ligation into the desired expression vector (pPpT4_αS, a modified version of the pPIC9 vector; Laroy & Contreras, 2000). In order to detect and purify the target proteins, they were equipped with a C-terminal hexahistidine tag. The engineered genes were codon-optimized and synthesized by GenScript.

The three synthetic sequences were delivered in a pUC57 vector (Amp^r) and were multiplied by transformation in *Escherichia coli* WK6 [(*lac-proAB*), *galE*, *strA*/F′ *lacI*^{q}, *lacZ*M15, *pro*A^+B^+; Zell & Fritz, 1987]. Using a GenePulser (Bio-Rad), electroporation was performed according to the manufacturer's instructions for bacteria: 1.8 kV, a capacitance of 25 µF and a resistance of 200 Ω. The transformed cells were selected on LB plates containing 100 µg ml^{-1} ampicillin and 2% glucose. Next, clones containing the insert [as confirmed by colony PCR using TvTS-generated primers: rTvTS1_FW (5′-ACTCGTGGATACTGGACCTT-3′), rTvTS3_FW (5′-TGCGAGAGGTTACTGGACTT-3′), rTvTS4_FW (5′-ACGTGGAAGACCGAAGTTATT-3′), rTvTS1_RV (5′-TGGCCAGAACAGTCCATGAT-3′), rTvTS3_RV (5′-AATGCACTCCGCCTGAATCG-3′) and rTvTS4_RV (5′-TTCGTCAGTCAGACGGATTC-3′)] were expanded overnight in 5 ml LB medium supplemented with 2% glucose and 100 µg ml^{-1} ampicillin (shaking at 200 rev min^{-1} at 310 K). The plasmids were extracted using a commercial kit (Invitrogen Plasmid miniPrep) according to the manufacturer's instructions. After digestion with Fermentas FastDigest AvrII and NotI restriction enzymes, the TvTS DNA inserts were incorporated by the use of Fermentas T4 DNA ligase into the linearized pPpT4α_S vector Zeo^r and introduced into *E. coli* DH5α electrocompetent cells (Invitrogen). The transformed cells were then selected on LB plates containing 50 µg ml^{-1} Zeocin (Invitrogen) according to the manufacturer's instructions. Colony PCR was performed using the commercially available AOX1 forward (5′-GACTGGTTCCAATTGACAAGC-3′) and reverse (5′-GGCAAATGGCATTCTGACATCCT-3′) primers from Invitrogen. pPpT4α_S clones containing the TvTS inserts were amplified overnight in LB medium supplemented with 2% glucose and 50 µg ml^{-1} Zeocin (shaking at 200 rev min^{-1} at 310 K) and the plasmids were extracted as described above. In order to confirm the correctness of the inserted sequences, the plasmids were sent for sequencing. Plasmids of each variant were further amplified and purified (as described above) in order to yield high amounts for transfection into the yeast *P. pastoris*. 10 µg of each pPpT4_αS_TvTS vector was digested with a threefold excess of New England Biolabs (NEB) PmeI restriction enzyme for 3 h in a water bath at 310 K. The linearized vector was purified according to the manufacturer's instructions using an NEB NucleoSpin Extract II Kit (eluted in 20 µl ultrapure water). Complete linearization of the vector was assessed by loading the digestion mixture onto an agarose gel.

Next, *P. pastoris* cells were made electrocompetent and transfected as described by Jacobs *et al.* (2009). A fresh *P. pastoris* colony was used to inoculate 5 ml yeast extract peptone dextrose [YPD; 1% yeast extract, 2% peptone, 2% dextrose (glucose)] medium and was grown overnight at 301 K with shaking at 250 rev min^{-1}. The next morning, the preculture was used to inoculate baffled shake flasks containing 250 ml YPD medium. The cells were allowed to grow overnight at 301 K with shaking at 250 rev min^{-1} until the culture reached an $OD_{600\,nm}$ of 1.5 (*i.e.* approximately 7.5 × 10^7 cells ml^{-1}). The cells were then harvested by centrifugation for 5 min [1500*g* (3000 rev min^{-1} in a F10S-6x500y rotor)] at 277 K. The supernatant was removed and the cells were resuspended in lithium acetate/DTT solution (100 m*M* lithium acetate, 10 m*M* DTT, 0.6 *M* sorbitol, 10 m*M* Tris–HCl pH 7.5) at 10^8 cells ml^{-1} and incubated for 30 min at 298 K with gentle shaking at 100 rev min^{-1}. The centrifugation step was repeated as described above, the supernatant was removed and the cells were resuspended in 1.5 ml ice-cold 1 *M* sorbitol per 5 × 10^8 cells. From this point on, the cells were kept on ice. The sorbitol wash described above was repeated two times. The cells were harvested by centrifugation for 5 min [1800*g* (2730 rev min^{-1} in a A-4-81 rotor)] at 277 K and the supernatant was removed. The cells were resuspended in 1 *M* ice-cold sorbitol to obtain a final concentration of 10^{10} cells ml^{-1}.

Fresh electrocompetent *P. pastoris* cells (80 µl) were then incubated with 300 ng linearized plasmid DNA for 5 min at 277 K in a pre-cooled Bio-Rad GenePulser electroporation cuvette (0.2 cm). Using a GenePulser (Bio-Rad), the cells were pulsed using the manufacturer's parameters for yeast: 1.5 kV, a capacitance of 25 µF, an extended capacitance of 125 µF and a resistance of 200 Ω. 1 ml ice-cold 1 *M* sorbitol was added directly after pulsing the cells. The transfected cells were transferred into a sterile 15 ml tube and incubated without shaking at 303 K for 1–2 h. Finally, 10, 50 and 200 µl of cell suspension were plated on YPD plates containing 100 µg ml^{-1} Zeocin and incubated for 2–3 d at 303 K until colonies appeared.

2.2. Protein production and purification

Small-scale expression clones containing the different rTvTS inserts from the YPD plates were inoculated in 2 ml buffered glycerol-complex medium (BMGY; 1% yeast extract, 2% peptone, 100 m*M* potasium phosphate pH 7, 1.34% yeast nitrogen base, 1% glycerol) in 24-well plates. After 24 h, the cells were collected and resuspended in 2 ml buffered methanol-complex medium (BMMY; 1% yeast extract, 2% peptone, 100 m*M* potassium phosphate pH 7, 1.34% yeast nitrogen base, 1% methanol). Methanol-induced expression was allowed for 48 h, with the addition of 0.5% methanol after 12 h. The plates were then centrifuged for 7 min [1600*g* (3120 rev min^{-1} in an A-2-DWP swing bucket rotor)] at 277 K. The collected yeast supernatants were screened for protein production using dot blotting. Detection of TvTS protein was performed using mouse anti-HIS (Sigma–Aldrich, catalogue No. H1029) followed by goat anti-mouse HRP (Sigma–

Aldrich, catalogue No. A4416). In order to optimize production yields, the constructs were transfected into three different *P. pastoris* strains (YA215, YA284 and YA550). Furthermore, the additives HisHAC or BHH were added to the culture medium to optimize growth and protein-production conditions. The Y215 strain (+ HisHAC) was selected as the best producer based on SDS–PAGE supernatant profiles.

For the reasons discussed in §3, large-scale production was only performed for rTvTS1. Here, 250 ml shaking flasks with 200 ml BMGY medium were inoculated with 1 ml cell precultures grown overnight in 5 ml BMGY in 50 ml Falcon tubes. After 48 h shaking at 250 rev min^{-1} and 303 K ($OD_{600\,nm}$ of around 15), cells were collected by centrifugation for 7 min [3000*g* (4200 rev min^{-1} in a F10S-6x500y rotor)] at 303 K and resuspended in 200 ml BMMY. Methanol-induced expression was allowed for 48 h at 250 rev min^{-1} and 303 K with addition of 1% methanol every 12 h. Cells were then collected by centrifugation for 7 min [3000*g* (4200 rev min^{-1} in a F10S-6x500y rotor)] at 277 K and the supernatant was collected. Finally, in order to proceed to the purification step, the supernatant was desalted by adding a 100-fold dilution of glutathione solution (200 m*M* $MgCl_2$ + 10 mg ml^{-1} glutathione), a 50-fold dilution of glutathione solution containing 1 *M* imidazole and adjusting the pH to 7.5 with 1 *M* NaOH. After precipitation of the salts, the supernatant was collected and filter-sterilized.

Ni–NTA chelating beads (HisPur Ni–NTA resin, Thermo Scientific) were used to capture and purify the recombinant protein from the yeast supernatant. Small-scale purification was performed in the same way for all TvTS variants. A PD-10 column (GE Healthcare) was packed with 2 ml Ni–NTA beads and equilibrated with phosphate-buffered saline (PBS) pH 7.0 over five column volumes. Next, the desalted supernatant was loaded onto the column followed by a wash with PBS pH 7.0 over ten column volumes. Finally, the rTvTS was eluted using PBS pH 7.5 containing 500 m*M* imidazole. All of the above steps were performed at room temperature, and gravitational force was employed to drain the solutions through the column. Large-scale purification was only performed for rTvTS1. A PD-20 column (GE Healthcare) was packed with Ni–NTA beads and connected to a peristaltic pump (Watson 101U/R). The column was extensively equilibrated with PBS pH 7.0 at 5 ml min^{-1}. The desalted supernatant was then passed through the column at 1.5 ml min^{-1}. The column was washed with PBS pH 7.0 containing 10 m*M* imidazole over two column volumes at 2 ml min^{-1}. Finally, the rTvTS was eluted using PBS pH 7.5 containing 250 m*M* imidazole. All of the above steps were performed at 277 K. For both small-scale and large-scale purification, the imidazole was removed from the eluate by overnight dialysis (Spectra/Por 1 MWCO 6–8 kDa) in 20 m*M* MES, 10 m*M* NaCl, 5% glycerol pH 6.0 at 277 K.

For the deglycosylation of rTvTS, 21 mg rTvTS1 in 20 m*M* MES pH 6.0, 10 m*M* NaCl, 5% glycerol was first concentrated to 7 mg ml^{-1} using a Vivaspin 20 MWCO (PES) 5000 and incubated overnight with 50 000 U endoglycosidase Hf (EndoHf; NEB) at 277 K on a table-top shaker at 100 rev min^{-1}. The sample was then centrifuged for 2 min [20 000*g* (14 000 rev min^{-1} in a F45-30-11 rotor)] at 277 K to pellet the carbohydrate residues. Using the predicted difference in pI between EndoHf and rTvTS1 (pI of 5.5 and 9.0, respectively), the EndoHf was separated from the deglycosylated rTvTS1 protein by ion exchange. After equilibration of the cation-exchange column (5 ml HiTrap SP FF, GE Healthcare) in 20 m*M* MES pH 6.0, 10 m*M* NaCl, 5% glycerol, the solution containing deglycosylated rTvTS1 and EndoHf was loaded onto the column. A NaCl gradient (from 10 m*M* to 1 *M*) was performed, in which the positively charged rTvTS1 protein eluted around 290 m*M* NaCl while the (close to) neutrally charged EndoHf eluted in the flowthrough. After purification, the deglycosylated product was concentrated (as described above) and the salt concentration of the buffer was decreased to 150 m*M*.

2.3. Activity assays

The following activity assays were performed at several stages throughout the production and purification steps to measure the biological activity of all produced and purified rTvTS constructs. Both sialidase and *trans*-sialidase activity assays were monitored *via* product release of the fluorophore 4-methylumbelliferone (MU). In both assays, enzyme activity was measured using a fluorometer (FluoStar Optima, BMG Labtech, Germany) with excitation at 365 nm and emission at 450 nm. Fluorescence was measured every minute for 2 h at 310 K. Blank values (empty-vector supernatant or buffer with substrate) were systematically subtracted. Fluorescence was calibrated as the amount of product formed by a previously defined MU standard curve. Recombinant *T. cruzi trans*-sialidase (rTcTS) was used as a positive control.

In the sialidase activity assay, 500 µ*M* 2′-(4-methylumbelliferyl)-α-D-*N*-acetylneuraminic acid (Mu-Neu5Ac) substrate and 0.2% bovine serum albumin were combined with the rTvTS protein sample in a final volume of 100 µl. Upon sialidase activity, the substrate is cleaved into α-D-*N*-acetylneuraminic acid (a sialic acid) and the MU fluorophore. In the *trans*-sialidase activity assay, 25 µl sample was incubated with 1 m*M* α2,3-sialyllactose (α2,3-SL), representing the donor, and 2 m*M* MU-α-D-galactopyranoside (MU-Gal), representing the acceptor, in a final volume of 50 µl at 310 K for 2 h. Upon TS activity, the sialyl group from the donor is transferred to the acceptor MU-Gal. The reaction was stopped by adding 1 ml ice-cold water. The solution was then incubated with 300 µl Q Sepharose Fast Flow resin (Sigma–Aldrich). The resin was activated prior to the experiment by overnight incubation in 1 *M* sodium acetate and washed three times with ice-cold water [centrifugation for 1 min at 2000*g* (3150 rev min^{-1} in an A-4-81 rotor) at 277 K]. The mixture was homogenized by inversion and centrifuged for 1 min [2000*g* (4340 rev min^{-1} in an F-45-30-11 rotor)] at 277 K and washed three times with ice-cold water. The negatively charged sialyl residues were thus captured by the anion-exchange Q Sepharose beads (both the remaining donor and the sialylated acceptor). The sialylated product (and remaining donor) was eluted by mixing the resin with 700 µl

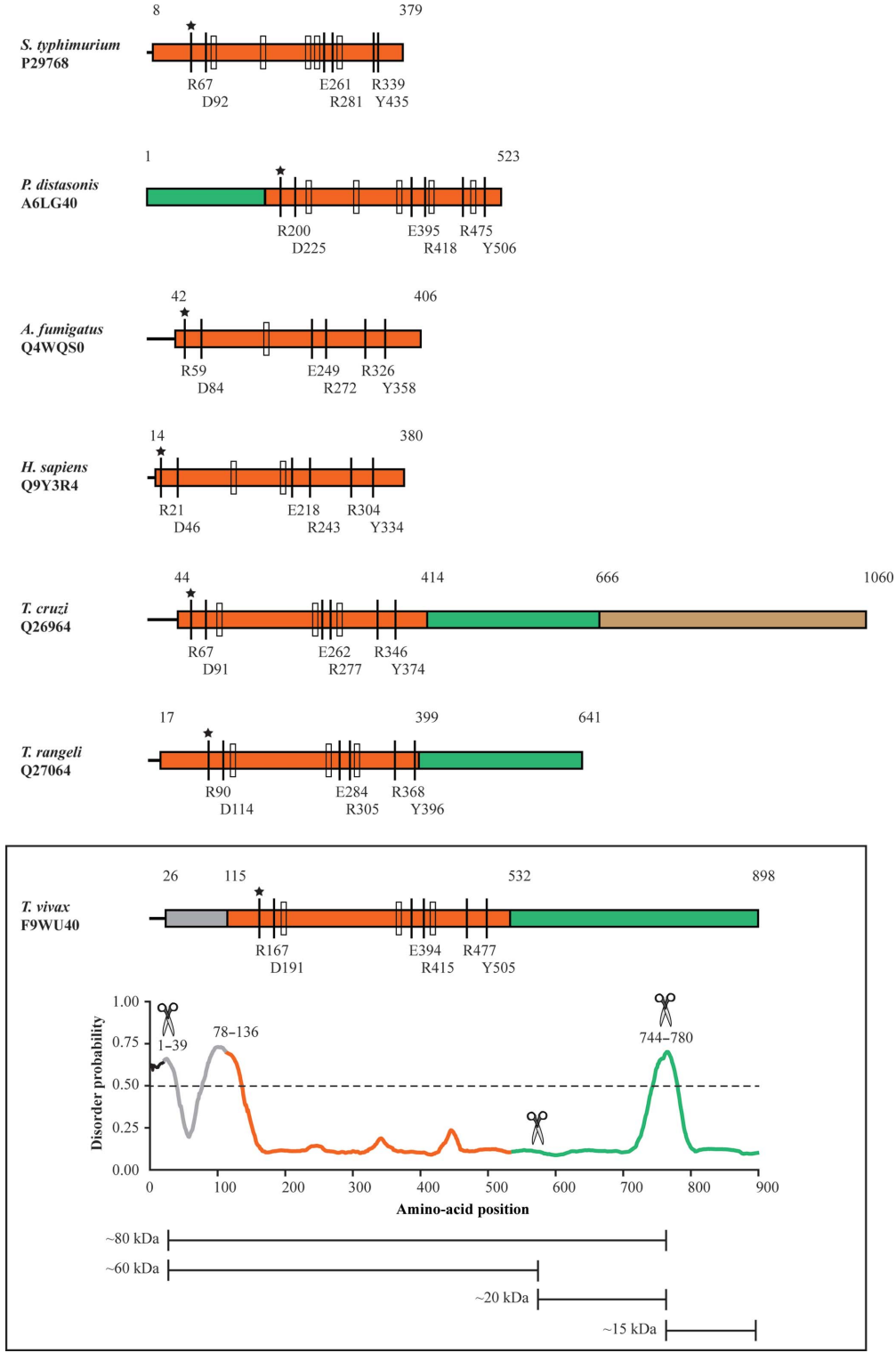


Figure 2
T. vivax TvTS1 contains the conserved 'sialidase motif'. The top panel displays a schematic representation of an alignment between sialidases and *trans*-sialidases from various organisms (from top to bottom): *S. typhimurium* sialidase (UniProt P29768), *P. distasonis* sialidase (UniProt A6LG40), *A. fumigatus* sialidase (UniProt Q4WQS0), *H. sapiens* sialidase (UniProt Q9Y3R4), *T. cruzi trans*-sialidase (UniProt Q26964), *T. rangeli* sialidase (UniProt Q27064) and *T. vivax* TvTS1 (UniProt F9WU40). All sequences contain the conserved sialidase residues (black bars) and the 'sialidase motif' (FR*x*P and S*x*D*x*G*x*T-W/*x*, indicated by a star and a rectangle, respectively). The catalytic domains, lectin-like domains and domains of unknown function are displayed as orange, green and grey boxes, respectively. The immunogenic SAPA repeats of *T. cruzi trans*-sialidase are indicated by the brown box. The alignment was generated using the *MAFFT* algorithm (Katoh *et al.*, 2002) within the *Geneious Pro* program suite (Biomatters Ltd). The bottom panel represents a *POODLE-L* (Hirose *et al.*, 2007) analysis of the TvTS1 sequence. Scores above 0.5 indicate disorder/flexibility and regions with predicted disorder are highlighted by their amino-acid numbers. The scissors indicate the proteolytic cuts generated during protein production and the four resulting products observed during SDS–PAGE are schematically depicted below the *POODLE-L* analysis.

1 *N* HCl. The eluate was collected by centrifugation for 2 min [2000*g* (4340 rev min^{-1} in an F-45-30-11 rotor)] at 277 K. In order to release the MU fluorophore from the sialylated acceptor for fluorescence measurements, the solution was hydrolysed at 368 K for 45 min and then allowed to thaw on ice. Finally, the sample was neutralized by first adding 290 µl 2 *N* NaOH and then 300 µl 1 *M* Tris–glycine pH 10.0.

2.4. Analytical gel filtration

Analytical gel-filtration experiments were performed on an ÄKTAexplorer platform (GE Healthcare) using a Superdex 200 HR 16/60 column (GE Healthcare) to determine the size and molecular mass of the deglycosylated rTvTS1. The column was pre-equilibrated using running buffer (20 m*M* MES pH 6.0, 150 m*M* NaCl, 5% glycerol) for at least one column volume. A 500 µl sample at a concentration of 6 mg ml^{-1} was injected using running buffer at a flow rate of 1.5 ml min^{-1}. To determine the molecular masses and the hydrodynamic radii of the proteins (Uversky, 1993), the column was calibrated using molecular-mass standards from Bio-Rad.

2.5. Crystallization, data collection and data processing

Crystallization conditions for purified, deglycosylated rTvTS1 were manually screened using the sitting-drop vapour-diffusion method in 96-well plates (MRC 2 well crystallization plates, Swissci) with drops consisting of 1 µl rTvTS1 (10 mg ml^{-1}) and 1 µl reservoir solution equilibrated against 100 µl reservoir solution. Commercial screens from Molecular Dimensions (ProPlex and PACT premier), Hampton Research (Grid Screen Salt HT) and Jena Bioscience (JBScreen Basic) were used for initial screening. All plates were incubated at 293 K. The C-terminal hexahistidine tag was retained for crystallization.

Of the 384 conditions screened, four produced promising hits (with the appearance of microcrystals) after incubation for 7–14 d. All hits were from the PACT premier screen: conditions A6 [100 m*M* SPG buffer pH 9.0, 25%(*w*/*v*) PEG 1500], A7 [200 m*M* sodium chloride, 100 m*M* sodium acetate pH 5.0, 20%(*w*/*v*) PEG 6000], A8 [200 m*M* ammonium chloride, 100 m*M* sodium acetate pH 5.0, 20%(*w*/*v*) PEG 6000] and B9 [200 m*M* lithium chloride, 100 m*M* MES pH 6.0, 20%(*w*/*v*) PEG 6000]. Optimization of these conditions was carried out through variation of the precipitant concentrations (8–22% PEG 1500 and PEG 6000), protein concentration (5–10 mg ml^{-1}) and salt concentrations (100 and 200 m*M*). Diffraction-quality crystals were obtained from an optimized condition A8 from PACT premier [200 m*M* sodium chloride, 100 m*M* sodium acetate pH 5.0, 16%(*w*/*v*) PEG 6000] at a protein concentration of 10 mg ml^{-1} after incubation for 5–7 d at 293 K.

The rTvTS1 crystals were cryocooled in liquid nitrogen with the addition of 15%(*v*/*v*) glycerol to the mother liquor as a cryoprotectant. A data set was collected from a single rTvTS1 crystal on the PROXIMA1 beamline at the SOLEIL synchrotron, Gif-Sur-Yvette, France. 360° of data were collected from the crystal in 0.2° slices using a Pilatus 6M detector with an exposure time of 0.2 s per degree, a crystal-to-detector distance of 392.60 mm and a wavelength of 0.98011 Å. The collected data set was processed with *XDS* (Kabsch, 2010) and the quality of the collected data set was verified by close inspection of the *XDS* output files and by using *phenix.xtriage* (Zwart *et al.*, 2005) in the *PHENIX* package (Adams *et al.*, 2010). Twinning tests were performed by *phenix.xtriage*. Analysis of the unit-cell contents was performed with *MATTHEWS_COEF*, which is part of the *CCP*4 package (Winn *et al.*, 2011).

3. Results and discussion

An *in silico BLAST* search of the *T. vivax* Y486 genome against the sequences of *T. rangeli* sialidase (UniProt code O44049) and *T. cruzi trans*-sialidase (UniProt code Q26966) resulted in the identification of five putative *trans*-sialidase sequences, which will be referred to as TvTS1, TvTS2, TvTS3, TvTS4 and TvTS5 (UniProt codes F9WU40, F9WQB0, F9WQX6, G0TZP8 and F9WTR7, respectively) in the following. As can be seen in Fig. 2 (top panel), these putative TvTSs contain the conserved residues which are the hallmarks of sialidases and *trans*-sialidases. In order to confirm the presence of TvTS1–TvTS5 at the DNA and mRNA levels, PCR was carried out on *T. vivax* lysates extracted from infected mice blood. While this experiment confirmed the existence of the TvTS genes in the genome, TvTS5 could not be detected at the transcriptional level. The existence of the remaining TvTSs at the protein level was checked by screening the *T. vivax* lysates and the serum of infected mice for sialidase/*trans*-sialidase activity. Given that the lysates and infected serum samples scored positive for sialidase/*trans*-sialidase activity in comparison to non-infected blood samples, we reasoned that at least one TvTS manifested itself as an active protein. Based on these results, TvTS1, TvTS3 and TvTS4 were selected for recombinant production. TvTS2 was omitted from this selection because of its high active-site residue identity with TvTS3. Despite optimization efforts during small-scale protein-production tests (different strains and different additives), only recombinant TvTS1 (rTvTS1) was produced in milligram quantities per litre of *P. pastoris* culture. Nevertheless, the three proteins (rTvTS1, rTvTS3 and rTvTS4) were purified by IMAC *via* their C-terminal hexahistidine tags in a small-scale purification setup. Only rTvTS1 was found to be a stable, soluble protein, while rTvTS3 and rTvTS4 precipitated shortly after purification.

Large-scale protein production and purification was only performed for rTvTS1. The enzyme was successfully produced and secreted by *P. pastoris* cultures and can be obtained with high purity using IMAC (Figs. 3*a* and 3*b*). After the purification procedure, typical yields varied between 4 and 5 mg pure protein per litre of *P. pastoris* culture. Interestingly, during its production and purification the rTvTS1 protein does not appear as a single band on SDS–PAGE. Instead, a pattern of four distinct bands consistently appears, suggesting that the protein could have been nicked or trimmed by proteases during its production. *In silico* flexibility predictions indeed

show that rTvTS1 most probably contains long loop regions which could be vulnerable to proteolytic degradation (Fig. 2, bottom panel). If rTvTS1 is proteolytically processed during its production, this could have a significant influence on the activity of the enzyme. Therefore, the purified (and possibly truncated) rTvTS1 was tested in both sialidase and *trans*-sialidase activity assays. From Fig. 3(*c*), it is clear that the enzyme retains its *trans*-sialidase and residual sialidase activities. Because of its production in *P. pastoris* and the occurrence of nine potential glycosylation sites, rTvTS1 is expected to be glycosylated. Because heterogeneity in the glycosylation pattern of rTvTS1 could hamper crystallization, the enzyme was deglycosylated by treatment with the endoglycosidase enzyme EndoHf. A comparison *via* SDS–PAGE of rTvTS1 before and after EndoHf processing clearly demonstrates that rTvTS1 was glycosylated during production in *P. pastoris*, as the four-banded pattern now runs with a lower molecular mass (Fig. 3*b*). While the residual sialidase activity of rTvTS1 is moderately affected by the deglycosylation, the *trans*-sialidase activity of the enzyme remains unaltered (Fig. 3*c*). Finally, to assess the behaviour of deglycosylated rTvTS1 in solution, analytical gel filtration was employed (Fig. 3*d*). The deglycosylated protein elutes as a single peak at a position expected for a globular protein of its size (an experimentally determined molecular mass of 93.9 kDa *versus* a theoretical molecular mass of 97.6 kDa and an experimentally determined R_h of 40.4 Å *versus* a theoretical R_h of 37.6 Å). This again strengthens the idea that loop regions of the protein are probably proteolytically trimmed during its production, without affecting its general three-dimensional structure and

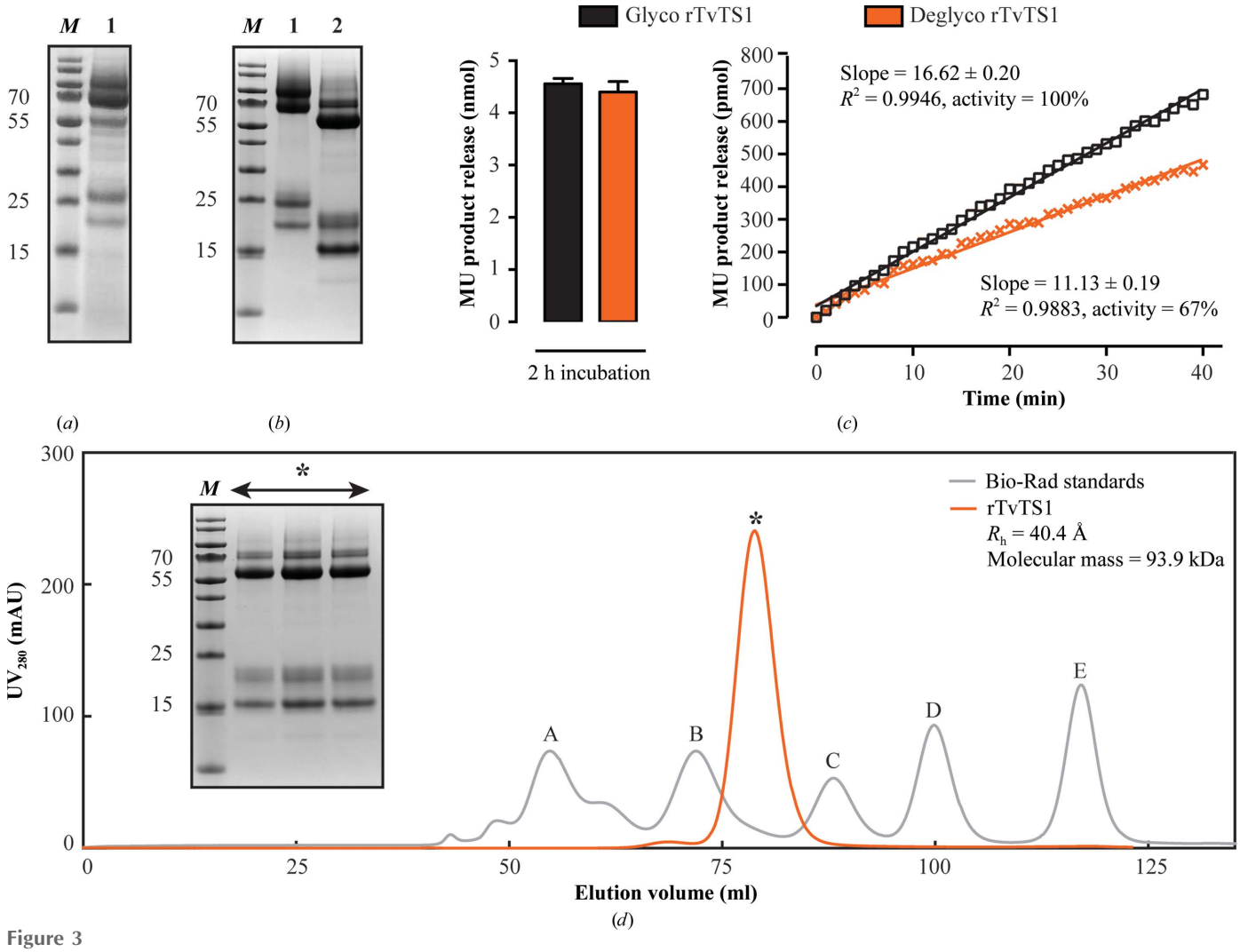


Figure 3
rTvTS1 is produced and purified as an active enzyme. (*a*) SDS–PAGE analysis of rTvTS1 production in *P. pastoris*. Lane 1, sample from the supernatant of an induced culture. Lane *M*, Prestained Protein Molecular Weight Marker (Fermentas; labelled in kDa). (*b*) SDS–PAGE analysis of glycosylated rTvTS1 purified *via* IMAC (lane 1). Deglycosylated rTvTS1 is shown in lane 2. Lane *M*, Prestained Protein Molecular Weight Marker (labelled in kDa). (*c*) *Trans*-sialidase (left panel) and sialidase (right panel) activity assay performed with 25 pmol glycosylated (black) and deglycosylated (orange) rTvTS1. (*d*) Analytical gel filtration. The experiment was performed on a Superdex 200 16/60 column. The orange and grey traces indicate the elution profiles for deglycosylated rTvTS1 and the Bio-Rad standards (A, bovine thyroglobulin, molecular mass 670 kDa, R_h = 86.0 Å; B, bovine γ-globulin, molecular mass 158 kDa, R_h = 51.0 Å; C, chicken ovalbumin, molecular mass 44 kDa, R_h = 28.0 Å; D, horse myoglobin, molecular mass 17 kDa, R_h = 19.0 Å; E, vitamin B_{12}, molecular mass 1.35 kDa), respectively. The inset shows an SDS–PAGE analysis of the elution peak. Lane *M*, Prestained Protein Molecular Weight Marker (labelled in kDa); the lanes marked with the asterisk are fractions from the deglycosylated rTvTS1 elution peak.

biochemical activity. This suspicion is confirmed by N-terminal sequencing of the four distinct bands observed during SDS–PAGE analysis. The N-terminal sequencing results demonstrate that the proteolytic cuts are localized in parts of the sequence which are expected to display increased flexibility based on *POODLE-L* analysis (Hirose *et al.*, 2007) or to be loop regions based on sequence alignment (Fig. 2).

Having obtained high-quality preparations of enzymatically active, deglycosylated rTvTS1, its crystallization was attempted. Crystallization conditions were manually screened by the sitting-drop vapour-diffusion method using various commercial screens. Of the conditions screened, four produced promising hits (the appearance of microcrystals) after incubation for 7–14 d. All hits were from the PACT premier screen: conditions A6 [100 m*M* SPG buffer pH 9.0, 25%(*w*/*v*) PEG 1500], A7 [200 m*M* sodium chloride, 100 m*M* sodium acetate pH 5.0, 20%(*w*/*v*) PEG 6000], A8 [200 m*M* ammonium chloride, 100 m*M* sodium acetate pH 5.0, 20%(*w*/*v*) PEG 6000] and B9 [200 m*M* lithium chloride, 100 m*M* MES pH 6.0, 20%(*w*/*v*) PEG 6000]. Optimization of these conditions was carried out through variation of the precipitant concentration (8–22% PEG 1500 and PEG 6000), the protein concentration (5–10 mg ml^{-1}) and the concentration of the salts present (100 and 200 m*M*). Diffraction-quality crystals were obtained using an optimized condition consisting of 200 m*M* sodium chloride, 100 m*M* sodium acetate pH 5.0, 16%(*w*/*v*) PEG 6000 at a protein concentration of 10 mg ml^{-1} after incubation for 5–7 d at 293 K (Fig. 4). To obtain single crystals that could be used for X-ray diffraction experiments, single needle crystals were dislodged from the crystal clusters (Fig. 4) using a cryoloop. The single rTvTS1 crystals were cryocooled in liquid nitrogen with the addition of 15%(*v*/*v*) glycerol to the mother liquor as a cryoprotectant and diffracted to a resolution of ∼2.5 Å (Fig. 4). The crystals belonged to space group $P2_12_12_1$, with unit-cell parameters a = 57.3, b = 78.4, c = 209.0 Å, and contained one rTvTS1 monomer in the asymmetric unit according to Matthews analysis (P = 1.0, 48.9% solvent content, V_M = 2.41 Å^3 Da^{-1}). The resulting data-collection statistics are given in Table 1.

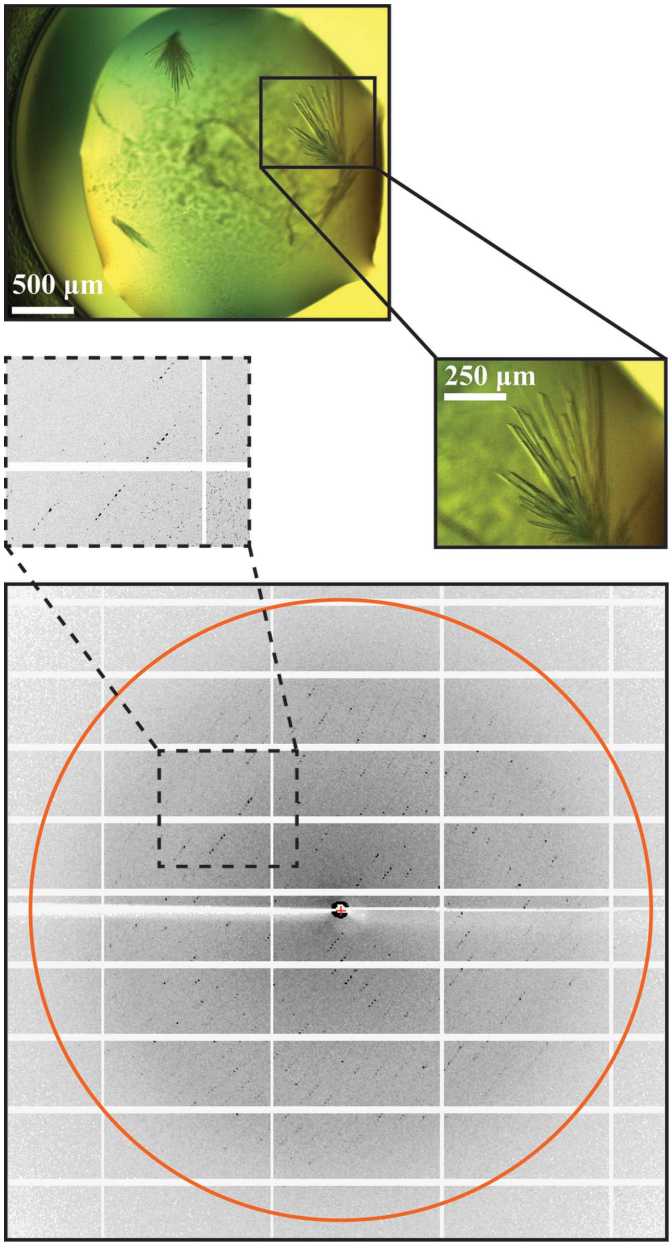


Figure 4
Morphology and diffraction of the rTvTS1 crystals. Top panel, rTvTS1 crystals obtained by the sitting-drop vapour-diffusion method (the crystallization condition is described in the text) appear as clusters. The black box displays an enlargement of the crystals. Bottom panel, a single rTvTS1 crystal diffracted to a resolution of around 2.5 Å (the 2.5 Å resolution limit is indicated by an orange circle). The dashed box displays an enlargement of the diffraction pattern.

Table 1
Data-collection statistics for rTvTS1.

Values in parentheses are for the highest resolution shell.

Beamline	PROXIMA1, SOLEIL
Wavelength (Å)	0.98011
Resolution range (Å)	44.26–2.48 (2.57–2.48)
Unit-cell parameters (Å, °)	a = 57.3, b = 78.4, c = 209.0, $\alpha = \beta = \gamma$ = 90
Space group	$P2_12_12_1$
Mosaicity (°)	0.247
R_{meas}†	0.15 (1.0)
No. of measured reflections	439446 (37268)
No. of unique reflections	33882 (3105)
Multiplicity	12.9 (12.0)
$\langle I/\sigma(I)\rangle$	14.721 (2.57)
Completeness (%)	99.32 (95.80)

† $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$ (Diederichs & Karplus, 1997).

The elucidation of the crystal structure of rTvTS1 will be interesting for several reasons. Firstly, the five newly identified *T. vivax trans*-sialidases display low sequence identities to other known *trans*-sialidases, including the well studied TcTS enzyme from *T. cruzi* (Buschiazzo *et al.*, 2012). While the sequence identities between TvTS1 to TvTS5 range between 51 and 69%, the sequence identity of TvTS1 to other trypanosomal (*trans*-)sialidases is much lower (between 19 and 32%). In addition, it seems that the TvTS enzymes, including TvTS1, contain an extra 130 residues at the N-terminal side of the *trans*-sialidase domain. The three-dimensional

architecture of this additional region as well as its function remain enigmatic. It is tempting to speculate that the differences between the *trans*-sialidases encoded by intracellular and extracellular trypanosomes might be related to the distinct functions that these enzymes play in the divergent life cycles of these two classes of parasite. Indeed, for the intracellular trypanosome *T. cruzi* it has been shown that TcTS is essential for host immune evasion and cell invasion by the parasite (Lieke *et al.*, 2011; Giorgi & de Lederkremer, 2011). The biological role of TvTS1 is less well defined. Evidence presented by Guegan *et al.* (2013) suggests that TvTS1 may play a role in *T. vivax*-associated anaemia, although this hypothesis requires further investigation. Study of the structure–function relationship of TvTS1 will help in grasping its biological function during bloodstream *T. vivax* parasite infections and might serve as the basis for the development of an antipathology treatment.

Acknowledgements

This work was supported by the IWT and by equipment grant UABR/09/005 from the Hercules Foundation. CLFH is an IWT PhD fellow (IWT sb-number 91466). We acknowledge BiostructX support for synchrotron access under project 6601. We are grateful to the staff of the PROXIMA1 beamline at SOLEIL, Gif-sur-Yvette, France for support with data collection and processing.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Amaya, M. F., Buschiazzo, A., Nguyen, T. & Alzari, P. M. (2003). *J. Mol. Biol.* **325**, 773–784.
Amaya, M. F., Watts, A. G., Damager, I., Wehenkel, A., Nguyen, T., Buschiazzo, A., Paris, G., Frasch, A. C., Withers, S. G. & Alzari, P. M. (2004). *Structure*, **12**, 775–784.
Barrett, M. P. & Fairlamb, A. H. (1999). *Parasitol. Today*, **15**, 136–140.
Blom-Potar, M. C., Chamond, N., Cosson, A., Jouvion, G., Droin-Bergère, S., Huerre, M. & Minoprio, P. (2010). *PLoS Negl. Trop. Dis.* **4**, e793.
Bratosin, D., Mazurier, J., Tissier, J. P., Estaquier, J., Huart, J. J., Ameisen, J. C., Aminoff, D. & Montreuil, J. (1998). *Biochimie*, **80**, 173–195.
Buratai, L. B., Nok, A. J., Ibrahim, S., Umar, I. A. & Esievo, K. A. N. (2006). *Cell. Biochem. Funct.* **24**, 71–77.
Buschiazzo, A., Amaya, F., Cremona, L., Frasch, A. C. & Alzari, P. M. (2002). *Mol. Cell*, **10**, 757–768.
Buschiazzo, A., Muiá, R., Larrieux, N., Pitcovsky, T., Mucci, J. & Campetella, O. (2012). *PLoS Pathog.* **8**, e1002474.
Chen, G. Y. & Nuñez, G. (2010). *Nature Rev. Immunol.* **10**, 826–837.
Chen, X. & Varki, A. (2010). *ACS Chem. Biol.* **5**, 163–176.
Cheng, J., Huang, S., Yu, H., Li, Y., Lau, K. & Chen, X. (2010). *Glycobiology*, **20**, 260–268.
Cheng, J., Yu, H., Lau, K., Huang, S., Chokhawala, H., Li, Y., Tiwari, V. K. & Chen, X. (2008). *Glycobiology*, **18**, 686–697.
Cross, G. A. & Takle, G. B. (1993). *Annu. Rev. Microbiol.* **47**, 385–411.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Mol. Biol.* **4**, 269–275.
Engstler, M., Reuter, G. & Schauer, R. (1993). *Mol. Biochem. Parasitol.* **61**, 1–13.
Gbem, T. T., Waespy, M., Hesse, B., Dietz, F., Smith, J., Chechet, G. D., Nok, J. A. & Kelm, S. (2013). *PLoS Negl. Trop. Dis.* **7**, e2549.
Giorgi, M. E. & de Lederkremer, R. M. (2011). *Carbohydr. Res.* **346**, 1389–1393.
Guegan, F., Plazolles, N., Baltz, T. & Coustou, V. (2013). *Cell. Microbiol.* **15**, 1285–1303.
Gut, H., King, S. J. & Walsh, M. A. (2008). *FEBS Lett.* **582**, 3348–3352.
Harrison, J. A., Kartha, K. P. R., Turnbull, W. B., Scheuerl, S. L., Naismith, J. H. & Field, R. A. (2001). *Bioorg. Med. Chem. Lett.* **11**, 141–144.
Hirose, S., Shimizu, K., Kanai, S., Kuroda, Y. & Noguchi, T. (2007). *Bioinformatics*, **23**, 2046–2053.
Jacobs, T., Erdmann, H. & Fleischer, B. (2010). *Eur. J. Cell Biol.* **89**, 113–116.
Jacobs, P. P., Geysens, S., Vervecken, W., Contreras, R. & Callewaert, N. (2009). *Nature Protoc.* **4**, 58–70.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002). *Nucleic Acids Res.* **30**, 3059–3066.
La Greca, F. & Magez, S. (2011). *Hum. Vaccines*, **7**, 1225–1233.
Laroy, W. & Contreras, R. (2000). *Protein Expr. Purif.* **20**, 389–393.
Lieke, T., Gröbe, D., Blanchard, V., Grunow, D., Tauber, R., Zimmermann-Kordmann, M., Jacobs, T. & Reutter, W. (2011). *Glycoconj. J.* **28**, 31–37.
Montagna, G., Cremona, M. L., Paris, G., Amaya, M. F., Buschiazzo, A., Alzari, P. M. & Frasch, A. C. C. (2002). *Eur. J. Biochem.* **269**, 2941–2950.
Montagna, G. N., Donelson, J. E. & Frasch, A. C. C. (2006). *J. Biol. Chem.* **281**, 33949–33958.
Muiá, R. P., Yu, H., Prescher, J. A., Hellman, U., Chen, X., Bertozzi, C. R. & Campetella, O. (2010). *Glycobiology*, **20**, 833–842.
Newstead, S., Chien, C.-H., Taylor, M. & Taylor, G. (2004). *Acta Cryst.* D**60**, 2063–2066.
Nok, A. J. & Balogun, E. O. (2003). *J. Biochem.* **133**, 725–730.
Pshezhetsky, A. V. & Ashmarina, L. I. (2013). *Biochemistry*, **78**, 736–745.
Schauer, R. & Kamerling, J. P. (2011). *Chembiochem*, **12**, 2246–2264.
Schenkman, S., Jiang, M.-S., Hart, G. W. & Nussenzweig, V. (1991). *Cell*, **65**, 1117–1125.
Schenkman, S., Pontes de Carvalho, L. & Nussenzweig, V. (1992). *J. Exp. Med.* **175**, 567–575.
Stafford, G., Roy, S., Honma, K. & Sharma, A. (2012). *Mol. Oral Microbiol.* **27**, 11–22.
Taylor, K. A. (1998). *Int. J. Parasitol.* **28**, 219–240.
Taylor, K. A. & Mertens, B. (1999). *Mem. Inst. Oswaldo Cruz*, **94**, 239–244.
Uversky, V. N. (1993). *Biochemistry*, **32**, 13288–13298.
Varki, A. (2011). *Glycobiology*, **21**, 1121–1124.
Varki, A. & Gagneux, P. (2012). *Ann. N. Y. Acad. Sci.* **1253**, 16–36.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zell, R. & Fritz, H.-J. (1987). *EMBO J.* **6**, 1809–1815.
Zwart, P. H., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsletter Protein Crystallogr.* **43**, contribution 7.